SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: INSTITUT PASTEUR
(B) STREET: 28 RUE DU DOCTEUR ROUX
(C) CITY: PARIS CEDEX 15
(E) COUNTRY: FRANCE
(F) POSTAL CODE (ZIP): 75724

(ii) TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA LIBRARY. APPLICATION TO THE DETECTION OF MYCOBACTERIA.

(iii) NUMBER OF SEQUENCES: 5

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12732 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ACCTGCGCTT GCAGAGATCA AATAGGGCGC ATGGGTCAGC ATAGTACAGG TCGTCGCGCA 60

TCTTTGATGC ATCGGAATAA GATGTCAGGC AATTAAAAGA GAAGCCACGG CGACTCGCGG 120

CATTCAGCAT GTCGAGCGTC GCTTCGATGT GAGCGCACCA TTCCGTGTCC AACGATTTCA 180

GACGAACATT GAATATTCCA CTCGCGACGC TATAGTCCGC CTCCCGATCT ATGCGCGCCG 240

CGCAGATGAA GTCTGCGTTC GCCCGACCTT CGAAACGTAG TGCGGCCGCG CGCACCATTT 300

CGGGGGAGAC GTCGATGCCG GTGTAATCAG TTTTGAAGCC ACGCGCATCT AGGTAGTCCA 360

GTAGAGCCCC ATAGCCACAG CCTAGATCGT TGATCGAAAA TGGGTCCGCC GCATTGACAA 420

TGCGCACCAG CTGGTCAAAG CGCAACGCCT GCCCGGCTTC GCCGTTCCAA TCGACGCCGC 480

GCGGGTGCCG TGTGCTTCGA GTTTCGATGC GTAGTAACGG GCCACGTCAG CGAGCATGGT 540

CGTTGCGTCT TCCGCCATGA AGCTGCCTCA CGATTTGTGT GTGTGGGCGT CGGTGCGTGG 600

GTCCGAGACT ATACCTTCAA CAGTTGCATG CCGAGGCTGC GGCGGGCAAT GACCCAAAAA 660

CCCGCCGGCA CGGTTCGCCG AGCAAGGAAG CGTGGAGACG ATAGATAATT TCACTGGCGA 720

CAGTACCTCA AATAGTCCGG AGCCTCGGCT CCGACGTTAA AGAGCAGATC CAGAATCGAC 780

ACGGCGGGCT CGAACCCTCC CCACAATTGC TTATAATCGC GGTAGCCGTC ATAATCGAAC 840

CAAGTTACCC GGATGCTAAG TTCGTCGAAC ACGCGCTCAT CGACATACGA ACGGGCTGAG 900

GGGCCAGAGA CATATTCGGT CGCTGCGGCC TGTTGGCAGA GGTTGGCCAG TCTCTCGGTC 960

TTGCCGTCGG CTAATTCGTA GTCCCACGAA TTTGCCAGTC GCGTGCTGAT ACCGAGATAA 1020

CTGCAAATCG CATTCAATAG ACGCCTGTTG AGTAAGGAAA GATTCGTGTG CTGTTCTTCG 1080

AGGTAAATCG GCGCGAGCCA GTCAGCGATC TCCGCAAAAT GAGCGGCCGC GCTGTAGTTG 1140

AATTCTAGTG CCCGCCAGTG CGCTTTCGCC CAATCGGTGC CGTCGATCAG CGTCTCACGT 1200

ATCTTTTGAT GGAAACGTCC CTTCACCTGG ACGGGAACAG TTATCCACTG TAACCCCTGG 1260

CTCGTTTTGA TCCGATTTCT GTTTCGCCAA TCACGCTTGG TATATTGCAT GTCATCATAG 1320

ATGATGAATT CATCGACGAA TGCAATCAGG TCAAAATATC CTCGCCAAGG TATGTAATTT 1380

GATTGAACAA TCGCGACTTT CTTCAACGCG GTGTCTCCAA TTTAGAATAA CAAATACGTC 1440

GCGCCCGCGA CAGCTCCGCT GGAGCGAGTT CAAGCGATTC TGCGACATAT TCAATATGGT 1500

GCTCGGGAAG GCCAGGATGG GCCGCGACCC GGGGCGTCCG GTGCGCGATG AACGTCGCAT 1560

CGTCTCCTGT GAGATAATTG CATCCGATCA TATAGGGCTG GCTGCGGCTA GGTTGCTGGC 1620

AAAAAGATAT CGCGGCCGAT CCGTTTCTGG TTTTGTCTTG ATGATCAAAT CCGCTTCCGT 1680

TCACGAGATC GATTCCTGGT CTTCCCCCAG CGTCGCGATG TCGATAGGTG TCGCGCTTTG 1740

TTCGTACCCG CACTACGCGG CGGCGAGAAC CTCGCCACCG AATCGGGATT GGGGGGAGGA 1800

TACCACTCGG TCGAGGCCCG TCACCGGCCT TCTAGCGGGT TGACCATCAG TGTTTGCAGG 1860

GCCCTATCCC GGTATGGCGC ACCACGGGAT CGGCAGCGTT CCGGTTGCTG GCGTGGTACC 1920

TCGTTGTGGC GCCGTGGTCC ATGTCGATTG AGTGCGTGGA TCAGTGTAAA CCGTTGCGCG 1980

CCATGTTCTG TAGGCACTGG TTCGGGTTGT GGTTAGGCTG CACGGTTGGC AGGTTACCAA 2040

CCACTGAGCC CCTGGGCGGA TGTGAGCTCG GACTCCGCCT ATGGGGTGTA ATTTTGGCAG 2100

ATTGGGCCGG GTCCCCGTGG TGAGGACTCC TCAACCGGAT TGGGTAAGCA TGAGGTGGTG 2160

CTGGCAGCGG TGTCCTGGTC GCTCTCCCGA GTAGGCCCGT TGTGACTGTC ATGTGGGCGA 2220

GCGGGTTTGC GCGCGTAGGA GACGATGATT ACTACGCACG TGACCAACCA CAAGAACGGT 2280

GCCCATGTCA CCGTGGTGAA AACGAGTGGC GTGGTACCGA CTACCCCTTT GGCTCCCAGC 2340

TGTCCATAGA GCGGCACGTA GAACGGCTGG CCCGGGACCG CGACGTTGAC GATGCTCAGC 2400

GCCACGGCCA AACTCACGCA GACGCCGACC GCGCGGCGGC GGTCTCCATG GGCTGCGAGT 2460

TGGTCGAATA TCCCAGCACC AGGAGGCCCG TTGGGGTCTC GGGCTACCAG TGCAGCGATT 2520

GGCAAGACGA AAACGAGATA GTAGAAGGCG ACGTCCGCGG GGGAGAAGGT GGCGGTGGCG 2580

AGCAACACAA TCCCCACCAT GACAGGCGGG ATACGGCGTC CGAGCGCCAG CACGGCGACC 2640

ACGACTATGA CTAGGACAGC AAACCCGATC TGCGTTCGCG GACCAGTGAG GAAACCCTCT 2700

GGGATCTTGC CCGATTGATA GTTCTTGATG CTATCGGGGA TCAGCAGGAG TGCCTTGCCA 2760

AAGGACACGT TCCGCGGGTC TCGAAGCCCT CCGAACGAAC TATTGAACTT GATGATGCCG 2820

TGGATCGACT GTGCGATCGT CCCCGGGAAG CCTCGTGGCC ACAACAGAAA GGCTGCGATA 2880

TTGGACACCA CCACGCCGGT GATCCCGATA CCAGCCCACC GCCATTGTCG AGCCGCCAAC 2940

AACACCACGC CGAGAACGAC GAACTGCGGC TTTACCAGGA CGGCCAAGAT CACCGTGATG 3000

GTGGCGAGGC CCCACCGCTG TCGGGACAAC GCCACGAAGT AAGCCAGCGC GATCGGTACC 3060

ACGAACCCTG TCGAGTTGCC TCGATCGATG ACCCCCCACG CCGGGATGGC CGCGGCGCCC 3120

AGTGTCACGA AGATGACCAC TCGCTCCAGA CCACGTGCCC CCCGGGCCGC CCAGATGGCG 3180

GGAGATATGA CCGCCATCGT TAGGGCGACC AGGTAACAGA TCAGCCCCAA GCGCGGCGCA 3240

CCCAGCCAAT GGCTGGGTAG TCCGAAAATC GCATACGGTA TGCGGGCGGG GGCCCATGCA 3300

GCAACCGCGG TCGGCTGGTA ATCGGCGGGT AGCGAGATCA GGTAGTCCGC GGGATTGGGT 3360

TGAATCCCGG CGGCGGCGAC CATGGCGTAG TCGCTGAAGC AGTGCCGACC GATATTCATG 3420

CCCCAATCAA GCCAACAGTC CCCAGGGACT ACCAAAAGAG TGGAAAAGAC GTCGACCGCG 3480

TACCACTGAC TGAGGGCGTA CGCCGTCGCC GCCGAAATCA CCGACGCCAG CAGGATGGTG 3540

CCGAGCATGA GGGTGCGCTC GGATTGGGAG CCGATCGCCC AGAGCCGCTC CCGGCTCGCG 3600

GTCACGGCAC CGCGCAACAC CTCCGGGGGT CGCTTCATCT GGATTCTCCT CGGTTCTGCG 3660

CGAAACGGTA GCAGAGCGCC ATGGTTGCCA ACGCGGTCGC CGGGCAGTCT AGACCGGATC 3720

TTCCTCGTGG CAACCGACAA CAGGACGTCG TTGCCGAAAG GGCGCTGGGC ACCGACATCT 3780

AGGATGAACC CACAGCCACG CCCCGACGTT ATGCCATGGC GAAGAGCGAC CGGCAGGAGC 3840

GGGAACCCAG TGAAGCGAGC GCTCATCACC GGAATCACAG GACCGGACGG CTCGTATCTC 3900

GCTAAGCTCC CGCTGAAGGG ATATGTGGCC GCTGGTAGCC CGGCCGAGGT CTATTTCTGC 3960

TGGGCGACAC GGAATTATCG CGAATTGTAT GGGTTGCTCG CGGTCAACAG CATCTGGTTC 4020

AATCACGAAT CACCGCGTCA CGGCGAGACA TTCATGACTC GTAATCCTGC ACCATATCGC 4080

GGTCGGCAAC GAGGCGCTGA TCGATGCGCA GACGCTGATG CGCCGGCCCA CCCGGATAGG 4140

TATCAGTATT GGGGCGTTCC GGCCAGCGTA CGAGGCGTGA TCGACCGCGC AATGGGTGTT 4200

TGCGTTGAGT AATAATCTGA ACCGTGTGAA CGCATGCATG GATGGATTCC TTGCCCGTAT 4260

CCGCTCACAT GTTGATGCGC ACGCGCCAGA ATTGCGTTCA CTGTTCGATA CGATGGCGGC 4320

CGAGGCCCGA TTTGCACGCG ACTGGCTGTC CGAGGACCTC GCGCGGTTGC CTGTCGGTGC 4380

AGCATTGCTG GAAGTGGGCG GGGGGGTACT TCTGCTCAGC TGTCAACTGG CGGCGGAGGG 4440

ATTTGACATC ACCGCCATCG AGCCGACGGG TGAAGGTTTT GGCAAGTTCA GACAGCTTGG 4500

CGACATCGTG CTGGAATTGG CTGCAGCACG ACCCACCATC GCGCCATGCA AGGCGGAAGA 4560

CTTTATTTCC GAGAAGCGGT TCGACTTCGC CTTCTCGCTG AATGTGATGG AGCACATCGA 4620

CCTTCCGGAT GAGGCAGTCA GGCGGGTATC GGAAGTGCTG AAACCGGGGG CCAGTTACCA 4680

CTTCCTGTGC CCGAATTACG TATTCCCGTA CGAACCGCAT TTCAATATCC CAACATTCTT 4740

CACCAAAGAG CTGACATGCC GGGTGATGCG ACATCGCATC GAGGGCAATA CGGGCATGGA 4800

TGACCCGAAG GGAGTCTGGC GTTCGCTCAA CTGGATTACG GTTCCCAAGG TGAAACGCTT 4860

TGCGGCGAAG GATGCGACGC TGACCTTGCG CTTCCACCGT GCAATGTTGG TATGGATGCT 4920

GGAACGCGCG CTGACGGATA AGGAATTCGC TGGTCGCCGG GCACAATGGA TGGTCGCTGC 4980

TATTCGCTCG GCGGTGAAAT TGCGTGTGCA TCATCTGGCA GGCTATGTTC CCGCTACGCT 5040

GCAGCCCATC ATGGATGTGC GGCTAACGAA GAGGTAATGA CATGGCGCAA GCGACATCGG 5100

GCATTCGCGC GGCACTTTCG CAACCTGCTG TGTATGAGGC GTATCAGCGG ATTGCGGGCG 5160

CTAAAAGCGG GCTTGCGTGG ATCACAACCG ACCCCATCCA GTCGTTGCCA GGCATGCGTA 5220

CTCTCGACCT CGGTTGCTGG CCAGCGGTGA TACACAGCTC CCCGCCAGTG GACGTGACAT 5280

GTACGAGAGA CGGCATGAGC GCGGAATGTG CGACCGTGCC GTCGAGATGA CCGACGTCGG 5340

CGCTACGGCA GCCCCCACCG GACCTATCGC GCGGGGCAGC GTCGCTCGGG TCGGCGCGGC 5400

GACCGCGTTG GCCGTTGCCT GCGTCTACAC GGTCATCTAT CTGGCGGCCC GCGACCTACC 5460

CCCGGCTTGT TTTTCGATAT TCGCGGTGTT TTGGGGGGCG CTCGGCATTG CCACCGGCGC 5520

CACCCACGGC CTCCTGCAAG AAACGACCCG CGAGGTCCGC TGGGTGCGCT CCACCCAAAT 5580

AGTTGCGGGC CATCGTACCC ATCCGCTGCG GGTGGCCGGG ATGATTGGCA CCGTCGCGGC 5640

CGTCGTAATT GCGGGTAGCT CACCGCTGTG GAGCCGACAG CTATTCGTCG AGGGGCGCTG 5700

GCTGTCCGTG GGGCTACTCA GCGTTGGGGT GGCCGGGTTC TGCGCGCAGG CGACCCTGCT 5760

GGGCGCGCTG GCCGGCGTCG ACCGGTGGAC ACAGTACGGG TCACTGATGG TGACCGACGC 5820

GGTCATCCGG TTGGCGGTCG CCGCGGCAGC GGTTGTGATC GGATGGGGTC TGGCCGGGTA 5880

CTTGTGGGCC GCCACCGCGG GAGCGGTGGC GTGGCTGCTC ATGCTGATGG CCTCGCCCAC 5940

CGCGCGCAGC GCGGCCAGCC TGCTGACGCC CGGGGGAATC GCCACGTTCG TGCGCGGTGC 6000

CGCTCATTCG ATAACCGCCG CGGGTGCCAG CGCGATTCTG GTAATGGGTT TCCCAGTGTT 6060

GCTCAAAGTG ACCTCCGACC AGTTAGGGGC AAAGGGCGGA GCGGTCATCC TGGCTGTGAC 6120

CTTGACGCGT GCGCCGCTTC TGGTCCCACT GAGCGCGATG CAAGGCAACC TGATCGCGCA 6180

TTTCGTCGAC CGGCGCACCC AACGGCTTCG GGCGCTGATC GCACCGGCGC TGGTCGTCGG 6240

CGGCATCGGT GCGGTCGGGA TGTTGGCCGC AGGGCTTACC GGTCCCTGGT TGCTGCGTGT 6300

TGGATTCGGC CCCGACTACC AAACTGGCGG GGCGTTGCTG GCCTGGTTGA CGGCAGCGGC 6360

GGTAGCTATC GCCATGCTGA CGCTGACCGG CGCCGCCGCG GTCGCGGCCG CACTGCACCG 6420

GGCGTATTTG CTGGGCTGGG TCAGCGCGAC GGTGGCGTCG ACGCTGTTGC TGCTGCTGCC 6480

GATGCCGCTG GAGACGCGCA CCGTGATCGC GCTGTTGTTC GGTCCAACGG TGGGAATCGC 6540

CATCCATGTG GCCGCGTTGG CGCGGCGACC CGACTGATTT GTGCCCCAGG TCGACAAATC 6600

ACGCCGTCTC GTCAGTGAGC ACTCCGTCCT CGGGTCCGAT CCTTCCAGGA GACGTTGCAA 6660

CCTGATTTGG CTCAAATTGG TGCGCACCGA GGGTCGGGCA CATCGTAGGG TCGCAACAGT 6720

CACATGTGTC ACTGCACCGG GCGACACCCG ATGTCCCGGC TCTCAGCGAC AGCTGTCTGA 6780

CCTGTGGTTT TGTTCCCAAG TTGGTCGTGG CTGTGCGGGA TTGGAGGTGG CGTGGGGGTC 6840

GCGTCGTATG GATTCTCCTC CTCGGTTCCG CGCGAAACGG CCGCAGGCGC AATGGTCACC 6900

AACTTGGCCG CGGTGGAGTC TAGCCTCACA TTTTCCTGGT CGCCCCCGAC AACCAGGAGG 6960

TCGCTGCAGA ACGGGCGTTC CCTACCCACA TCTACTATGA AGCGACAGCG GCGCCCCGCT 7020

GTGATGGCTG AGCATGACCG ACAGAGGCGG GAAGACAGTG AAGCGAGCGC TCATCACCGG 7080

AATCACCGGC CAGGACGGCT CGTATCTCGC CGAACTGCTG CTGGCCAAGG GGTATGAGGT 7140

TCACGGGCTC ATCCGGCGCG CTTCGACGTT CAACACCTCG CGGATCGATC ACCTCTACGT 7200

CGACCCGCAC CAACCGGGCG CGCGGCTGTT TCTGCACTAT GGTGACCTGA TCGACGGAAC 7260

CCGGTTGGTG ACCCTGCTGA GCACCATCGA ACCCGACGAG GTGTACAACC TGGCGGCGCA 7320

GTCACACGTG CGGGTGAGCT TCGACGAACC CGTGCACACC GGTGACACCA CCGGCATGGG 7380

ATCCATGCGA CTGCTGGAAG CCGTTCGGCT CTCTCGGGTG CACTGCCGCT TCTATCAGGC 7440

GTCCTCGTCG GAGATGTTCG GCGCCTCGCC GCCACCGCAG AACGAGCTGA CGCCGTTCTA 7500

CCCGCGGTCA CCGTATGGCG CCGCCAAGGT CTATTCGTAC TGGGCGACCC GCAATTATCG 7560

CGAAGCGTAC GGATTGTTCG CCGTTAACGG CATCTTGTTC AATCACGAAT CACCGCGGCG 7620

CGGTGAGACG TTCGTGACCC GAAAGATCAC CAGGGCCGTG GCACGCATCA AGGCCGGTAT 7680

CCAGTCCGAG GTCTATATGG GCAATCTGGA TGCGGTCCGC GACTGGGGGT ACGCGCCCGA 7740

ATACGTCGAA GGCATGTGGC GGATGCTGCA GACCGACGAG CCCGACGACT TCGTTTTGGC 7800

GACCGGGCGC GGTTTCACCG TGCGTGAGTT CGCGCGGGCC GCGTTCGAGC ATGCCGGTTT 7860

GGACTGGCAG CAGTACGTGA AATTCGACCA ACGCTATCTG CGGCCCACCG AGGTGGATTC 7920

GCTGATCGGC GACGCGACCA AGGCTGCCGA ATTGCTGGGC TGGAGGGCTT CGGTGCACAC 7980

TGACGAGTTG GCTCGGATCA TGGTCGACGC GGACATGGCG GCGCTGGAGT GCGAAGGCAA 8040

GCCGTGGATC GACAAGCCGA TGATCGCCGG CCGGACATGA ACGCGCACAC CTCGGTCGGC 8100

CCGCTTGACC GCGCGGCCCG GGTCTACATC GCCGGGCATC GCGGCCTGGT CGGGTCCGCG 8160

CTGCTACGCA CGTTTGCGGG CGCGGGGTTC ACCAACCTGC TGGTGCGGTC ACGCGCCGAG 8220

CTTGATCTGA CGGATCGGGC CGCGACGTTC GACTTCGTTC TCGAGTCGAG GCCGCAGGTC 8280

GTCATCGACG CGGCGGCCCG GGTCGGCGGC ATCCTGGCCA ACGACACCTA CCCGGCCGAT 8340

TTCCTGTCGG AAAACCTCCA GATCCAGGTC AACCTGCTGG ATGCCGCCGT GGCGGCGCGG 8400

GTGCCGCGGC TGCTGTTCCT GGGCTCGTCG TGCATCTACC CGAAACTCGC CCCGCAGCCG 8460

ATCCCGGAGA GCGCGCTGCT CACCGGTCCG TTGGAGCCGA CCAACGACGC GTACGCGATC 8520

GCCAAAATCG CCGGCATCCT TGCGGTCCAG GCGGTGCGCC GCCAACATGG CCTGCCGTGG 8580

ATCTCGGCGA TGCCCACCAA CCTGTACGGG CCAGGCGACA ACTTTTCGCC GTCCGGCTCG 8640

CATCTGCTGC CGGCACTCAT CCGCCGCTAT GACGAGGCCA AAGCCAGTGG CGCGCCCAAC 8700

GTGACCAACT GGGGCACCGG CACGCCCCGA CGGGAGTTGC TGCACGTCGA CGACCTGGCG 8760

AGCGCATGCC TGTATCTGCT GGAACATTTC GACGGGCCGA CCCATGTCAA CGTGGGAACC 8820

GGCATCGACC ACACCATCGG CGAGATCGCC GAGATGGTCG CCTCGGCGGT AGGCTATAGC 8880

GGCGAAACCC GCTGGGATCC AAGCAAACCG GACGGAACAC CACGCAAACT GCTGGATGTT 8940

TCGGTGCTAC GGGAGGCGGG ATGGCGGCCT TCGATCGCGC TGCGCGACGG CATCGAGGCG 9000

ACGGTGGCGT GGTATCGCGA GCACGCGGGA ACGGTTCGGC AATGAGGCTG GCCCGTCGCG 9060

CTCGGAACAT CTTGCGTCGC AACGGCATCG AGGTGTCGCG CTACTTTGCC GAACTGGACT 9120

GGGAACGCAA TTTCTTGCGC CAACTGCAAT CGCATCGGGT CAGTGCCGTG CTCGATGTCG 9180

GGGCCAATTC GGGGCAGTAC GCCAGGGGTC TGCGCGGCGC GGGCTTCGCG GGCCGCATCG 9240

TCTCGTTCGA GCCGCTGCCC GGGCCCTTTG CCGTCTTGCA GCGCAGCGCC TCCACGGACC 9300

CGTTGTGGGA ATGCCGGCGC TGTGCGCTGG GCGATGTCGA TGGAACCATC TCGATCAACG 9360

TCGCCGGCAA CGAGGGCGCC AGCAGTTCCG TCTTGCCGAT GTTGAAACGA CATCAGGACG 9420

CCTTTCCACC AGCCAACTAC GTGGGCGCCC AACGGGTGCC GATACATCGA CTCGATTCCG 9480

TGGCTGCAGA CGTTCTGCGG CCCAACGATA TTGCGTTCTT GAAGATCGAC GTTCAAGGAT 9540

TCGAGAAGCA GGTGATCGCG GGTGGCGATT CAACGGTGCA CGACCGATGC GTCGGCATGC 9600

AGCTCGAGCT GTCTTTCCAG CCGTTGTACG AGGGTGGCAT GCTCATCCGC GAGGCGCTCG 9660

ATCTCGTGGA TTCGTTGGGC TTTACGCTCT CGGGATTGCA ACCCGGTTTC ACCGACCCCC 9720

GCAACGGTCG AATGCTGCAG GCCGATGGCA TCTTCTTCCG GGGCAGCGAT TGACGCGCCG 9780

GCGCGTCAAT CTATTTCGAC ATTCGCGTGA AGACGTTTTC CCAGAATCGA CTGTTGTAGG 9840

CGTAGAACTC CCGGCCGCGT AGGTAGGCAT GTGATATTCG CCTTCCCCCG AACGGGTAGC 9900

GGCGATGAAG GTCGCCCATG CGGCGCAGAT CACCGAAGAC CGCGCTTGGT TCCCGGTGCG 9960

AGCCGACGCC CGTGGTGTCG AACTCGCACA GCACACACCG AATCGTGACC GGCTCGCATA 10020

CCAGCGCGGC CCGCAATATG AATTCCTGGT CGGCGGCGAT CCCGAAATCA AGGTCGTAGC 10080

CACCGATCTT GGCCACCAGC GATGATCCGA AGAACGATGC TTGATGCGGA ACAACCTGCT 10140

TGCCGGCCAG GAATTTGCGC AGGCTGAAAG GTATCGGGCC GCGCACCCGA TCGAGCCCGA 10200

CGAGACGATC CATCCCGAAG CCCCACAATT CGGACACCGG TCCCTTGCCG GATAGCGCCT 10260

CCACGGCCTG GGCTACCACG TCGGGCCCGG AAAAACGATC GGCGGAGTGC AAGAACCACA 10320

ACAGATCACC CGATGCGTGC GCGATGCCCT GGTTCATCGC GTCGTACCGC CCGCCGTCGG 10380

GCTCGGACTG CCAATACGCG AAGCCTGGTT CACACCCGGA CAGGTATGCC ACCACGTCGT 10440

CGCCGCTGCC ACCGTCGATT ACGATGTGCT CGATGCGTCC CCGGTAGCGT TGCGCCCGCA 10500

CACTTTTCAC CGTGCGCTGC AACCCGTCGA GGTCGTTGAA CGAGATCGTT ATCACCGAGA 10560

CGGTCGGAGC AGACGTCACC GAGTTCCCCT AGGTTGCTGG CGGCGATTGT GGATCACCGG 10620

GTCTTGATAC CGATGAAGGT GCCTCGAAGA TTCGCCGCAT AGGAACCTCC GAGCAACGAC 10680

TCGGCGATGC TTGGTTCCAA GTTGTCGTAC TCCTCCATCA CCAGGTCGAC GCCGACGTCT 10740

TTGATGGCCT GAAGTAGGTG CTCGCGTTGA ATCCAGAATG ACCGGCGATT GTCCCAGGAC 10800

GCCCATTTTG CGGTGTCGCG CTGGCCAAAC GAGCGGTCGT CGGAAAACTC GGTAAACCAC 10860

CTACCGGGAA GTCCCTCATG TTCGGTGGGC GCCGAGAGCA TGAACTTCAC CGGCGCCGGC 10920

CGCCGCAGCA ACCGATCGGT CAATTGTCGT GCCGTCGTGG GCAACCGGAG CCATTTATCG 10980

CTCCGGTTGA TGATCGAGAA GTGCGTCTGG AGAATCAGCA GCTTGTTCGT TACCGACGAG 11040

AGGGTTTCCA GGTATTGCTT CGGATTCTCC AGGTGGTAGA AGAGGCCGCA GCAGAAGACG 11100

GTATCGAAGA GCCCGTGGTT GGCGATGTTG AGGGCGTTGT CGTGGACGAA CCGGAGATTC 11160

GGCAGGTTGG TCTTCGATTT GATGTAGTTG CAGGCCGCCA TGTTCAGCTC GCGAACCTCG 11220

ATCCCGAGGA CCTGAAATCC CATGCGCGCG AACCCGACCG CGTACCCGCC TTCCAAGCAG 11280

CCGACATCGG CCAGGCGTAG GTGGCTCTTG TCCCCGGGAA AGACGGTTTC CAGAATCCCG 11340

CGCGCCGAGA TGAACCAGGA CGATTCGTCT AACGTGCGCG AGGACTCCGG TATCGTCAAG 11400

GTTCCGTCGT CGAGGCGAAC GTTGTGGGCG GTGAATTGTA CCGCGCCGGC CGAATGTTCC 11460

TGTGCCATCA CTTGGTTAGC CCCTTCGGCT GGTCCTGGGT TTGTCGACAT GGTCAGGCTC 11520

GACAGCCGCG TCGGAGCCGG GAGGGCCACA CATCCACGAG CCCCCTGCGG CTCGGCGTCG 11580

CGGCGGCGAG CTTGCGCCAC TGGGTCTTGA GCCGCCGCGC GGGTGTCGCC CCGCGGTGCT 11640

GCAGCGCCAG CATGGCGATC CGGGGATGGC GCGCGATGGT TTCCTGCAGC GCGGCGCGCC 11700

CCTCCGGGCC TGGAACGTTG GCGATCTGGC GAAGGATCCA GTCGGCCATG ACGGCGATGA 11760

GCTCCTCGCG CGCGGGGTCT CCCGGGAACA GGTCGAGCAT CGCGTCAAAC GTCGCCGCAT 11820

GCCCCGGACC CTGCGTCAAC CAGAACTTTG GCGGGTCCAC CACCTGGTTG TGCCACATGC 11880

CTTGGGCGTG GCGGCGATAC ACGGCCATGG TGTCGGGCAA CATGGCGATG TCGCCATGCA 11940

CCGCGTGCCG GACGTGCAGA TACCAGTCCA GGGGCATGAC GTCGGCAGGA ATGTCGTCGT 12000

AGCGCTCGAG GCGACGGTAC ACGGCCGAGT TGGTCTGGAT GAAGTTCATC AAGATCAACG 12060

CATCCAGGCT CAAGTTGCCC CGCACCCGAA CCGGGGGGAA CTTCGAGTCC TTGGCATGGC 12120

CGTCCTCCCA TATCACTCGG ACGGGATGGA AGCACACCGT CGTCTTGGGG TGCCGGTCGA 12180

GGAATGCGAC CTGTTTGCTT AGCTTCAGCG GATCGATCCA GTAGTCGTCC GCCTCGCACA 12240

ACGCGACGTA CTCGCCGCGA GCGGCCGACA GGGCGCCGGT CAGGTTCCCA TTGAGGCCGA 12300

GGTTTTCGGT CCTGAAGATC GGCCGGAACA CGTGCGGGTA CCGCTCGGCG TACTCACGGA 12360

TGATCGCCGG GGTGGCATCG GTCGACGCGT CGTCGGCGAC GATGATCTCC ACCGGGAAGT 12420

CGGTTTGCTG GTCGAGAAAG CTGTCGAAGG CCTGACGGGC GTAGCCCGCC TGGTTGTGAG 12480

TGGTCGAGAC GATGCTCACC TTGGGGCAAA GCTGGGGACT CACCGTCGGC CCTTTTCCTG 12540

CGCGGCCGCA AGGGTATTGC GATGGCGAAC GTGAATCGCC TGTGCCCGCC GGCCGTCGGC 12600

CGTCGTGGCC TGGTGGTCGG CGGACGTACG GCACACGCTG GCGAAGTATA GCGAGGGTGC 12660

ACTGACGTTG GGCTCGAACC GCGTGGCGCG CGGTGTGGGC GCACCGTCTC GAGTCGGTGC 12720

TGGTTGGCTC GC 12732

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 289 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ATACTCAAGC TTGCCGCAAT CGAAACCAAC CTGTTTGTGC CGCAAGAAAT TACGCCGTGG 60

CCCGGCGCCG ATCAAGAAAC GCCCCGGCGC GCGGCGGTGT CGTCGTATGG CATGACGGGC 120

ACCAATGTGC ACGCCATTGT CGAGCAGGCA CCGGTGCCAG CCCCCGAATC CGGTGCACCA 180

GGCGACACCC CGGCCACACC CGGTATCGAC GGCGCGCTGC TGTTCGCGCT GTCGGCCAGC 240

TCGCAGGACG CGCTGCGGCA AACCGCCGCG CGGCTGGCCG ATTGGGTCT 289

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 278 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TTGGCGGGTT GGCCACACAC CCGCCGGTGA CGGCGACGAT GCTGGGCTGG TTGCGGCCCT 60

GCGCCACCGC GGCTTGCATG CTGGTTGGCT GTCTTGGGAC GATCCCGAAA TAGTCCACGC 120

GGATCTGGTG ATTTTGCGGG CTACCCGCGA TTACCCCGCG CGGCTCGACG AGTTTTTGGC 180

CTGGACTACC CGCGTGGCCA ATCTGCTGAA CTCGCGGCCG GTGGTGGCCT GGAATGTCCA 240

CGCCGTTCAC CTACGTGACC TTGATGGGAT CCGGGGGT 278

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1280 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CCGACCCAGA CACTGACCGG GCGACCGCTG ATCGGCAACG GCACCCCCGG GGCGGTCGGC 60

AGCGGGGCCA CCGGGGCCCC CGGTGGGTGG CTGCTCGGCG ACGGCGGGGC CGGCGGGTCC 120

GGCGCGGCGG GCTCGGGCGC GCCCGGCGGG GCGGGCGGGG CTGCCGGGCT GTGGGGTACC 180

GGCGGGGCCG GCGGGATCGG CGGAGCCAGC ACCGTACTCG GCGGCACCGG CGGGGGAGGC 240

GGGGTCGGTG GGCTGTGGGG CGCCGGTGGG GCCGGCGGGG CCGGTGGAAC CGGCCTTGTT 300

GGTGGCGACG GCGGGGCCGG TGGGGCCGGC GGGACCGGCG GACTGCTGGC CGGGCTGATC 360

GGTGCCGGCG GAGGTCACGG CGGGACCGGC GGGCTCAGCA CTAATGGCGA CGGCGGGGTT 420

GGCGGGGCCG GCGGGAATGC CGGAATGCTC GCCGGGCCGG GCGGCGCCGG CGGAGCCGGC 480

GGTGACGGCG AAAACCTGGA CACCGGTGGG GACGGCGGGG CCGGCGGTAG CGCAGGGCTG 540

CTGTTCGGCA GCGGCGGCGC CGGCGGCGCC GGCGGATTTG GTTTCCTCGG TGGGGACGGC 600

GGGGCCGGTG GCAACGCCGG GCTGCTGTTG TCCAGCGGCG GGGCCGGCGG GTTCGGCGGG 660

TTCGGCACCG CCGGTGGGGT CGGTGGGGCC GGCGGCAATG CCGGCTGGCT GGGCTTCGGC 720

GGGGCCGGGG GCATCGGCGG AATCGGCGGT AACGCTAACG GGGGCGCCGG TGGGAACGGC 780

GGCACCGGCG GTCAGTTATG GGGTAGCGGC GGCGCCGGCG TCGAAGGCGG CGCAGCCTTA 840

AGCGTCGGCG ACACCGGCGG GGCCGGTGGC GTCGGCGGCA GCGCCGGGCT GATCGGCACC 900

GGCGGCAACG GCGGCAACGG CGGCACCGGC GCCAACGCCG GCAGCCCCGG AACCGGCGGC 960

GCCGGCGGGT TGCTGCTGGG CCAAAACGGG CTCAACGGGT TGCCGTAGCC GGGCGGCACG 1020

GCATGGCTTC CGGGCGTCAA CCACTCGCCG GTGATGCAGA TCGGCTGCGG AGCGGGCCGC 1080

CAAAATGGGG GCCGCCGCGC CAGGTATCTC GGCGAAGATC CCCGGCGCTC GAGCGCTTTG 1140

TCAGAGGCCC GTCGCGGGTC GTCGTGACGA CGGCTATCCG GGCGGTGCGG GTTTCGCGGC 1200

GCGCCCTGTG CCCGGCACCG CCGCCCGTTT GTCGGCAACG CCGCCGCGAC CCGTGAGCCG 1260

TCCAGCAGCT GGCGCCTGCG 1280

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 127 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGGCATCGGC GGAATCGGCG GTAACGCTAA CGGGGGCGCC GGTGGGAACG GCGGCACCGG 60

CGGTCAGTTA TGGGGTAGCG GCGGCGCCGG CGTCGAAGGC GGCGCAGCCT TAAGCGTCGG 120

CGACACC 127